Canadian Political Science Review Vol. 9, No. 1, 2015, 41-54

The Accuracy of Public Polls in Provincial Elections

David Coletto
Arthur Kroeger College of Public Affairs, Carleton University - Email address: dcoletto@carleton.ca

Bryan Breguet
Founder, TooCloseToCall.ca - Email address: bryanbreguet@hotmail.com

Abstract : This study extends work on the accuracy of polls released in the final week of election campaigns in Canada, using data from the nine provincial elections held in Canada between 2011 and 2013 to identify what might affect the accuracy of those polls. Specifically, we attempt to empirically test two arguments - mainly that some methodologies are better than others in measuring voter preferences and that the nature of the election itself might make accurate forecasts more difficult. We find that absolute change in voter turnout was the strongest predictor of polling accuracy. In contrast, sample size, survey mode, or electoral volatility were statistically significant predictors of polling accuracy.

Keywords: polling, polls, voter preference, electoral behaviour

Résumé: Cette étude poursuit les travaux effectués sur la précision des sondages publiés au cours de la dernière semaine des campagnes électorales au Canada. Cette étude utilise les données sur les neuf élections provinciales tenues au Canada entre 2011 et 2013 dans le but d'identifier les facteurs susceptibles d'intervenir dans la précision de ces sondages. Notamment, nous tentons de valider de façon empirique deux hypothèses – l'une selon laquelle certaines méthodes l'emporteraient sur d'autres pour mesurer les préférences des électeurs et l'autre voulant que la nature même de l'élection puisse faire en sorte qu'il soit plus difficile de faire des prévisions précises. Nous avons constaté qu'une variation absolue du niveau de participation électorale constitue l'indicateur le plus fiable de la précision des sondages. Par ailleurs, la taille de l'échantillon, le mode de sondage et la volatilité électorale se sont également avérés des facteurs prédictifs statistiquement significatifs de la précision des sondages.

Mots-clés: les sondages, la préférence des électeurs, le comportement électoral

Theoretical Framework

The provincial elections held between 2011 and 2013 had their share of publicly released polls. In total, at least 110 province-wide public polls were released for the nine elections held over the two year period. But along with the large number of public polls released during this time period, the provincial elections under scrutiny in this book are also remembered for how incorrect the polls were in forecasting three of the elections: Alberta, Quebec, and British Columbia.

Conducting a public opinion survey today is not as easy as it once was. In an age when every citizen had a landline, most answered their phone, and more often than not participated in a survey when they picked up the phone, conducting a poll was a relatively easy undertaking. Pull a random sample of telephone numbers, call the numbers and Canadians would answer a pollster's question.

But today, fewer Canadians have landlines, fewer Canadians are answering their phones, and even when someone answers the phone, few take the time to respond to the survey. Technological, cultural, and demographic changes have made public opinion polling more difficult. These challenges not only make it more difficult to generate a representative sample, but the cost to conduct the research has increased substantially. The emergence of alternative methodologies such as interactive voice response and internet data collection have lowered the cost of conducting the research, but raised additional questions about whether the sample of respondents they reach can be representative of the population.

Our intent is neither to ignore the problems faced by researchers during the Alberta, Quebec, and British Columbia elections nor to make excuses for the pollsters. For the sake of full disclosure, we should note that one of us works for a public opinion research firm, Abacus Data, and conducted publicly released polling during the Ontario, Alberta, British Columbia, and Nova Scotia provincial elections. Our objective is to try and understand what might cause polling in some elections to be more accurate than others. In this article, we consider polling accuracy to be how close final publicly released decided voter proportions by party are to the actual proportion of votes cast in an election.

During an election campaign, the stakes are high for public opinion researchers. The news media, pundits, and the public expect polls to accurately forecast the final election results. Moreover, the media often treat polls "as matters of fact with their limitations rarely discussed" (Andersen, 2000). Elections are one of the rare occasions the accuracy of public opinion research can be assessed and measured against a real outcome. However, rarely has there been a thorough study of how accurate pre-election polls are either at the federal or provincial level in Canada. Unlike in the "polling failures" in the United States in 1948 (Mitofsky, 1998) or Britain in 1992 (Jowell et al., 1993) forced the professional and academic research communities to assess polling methodologies and standards, Canada has not faced such a crisis.

The failure of polls to accurately forecast the results of the Alberta and British Columbia elections give us reason to ask why polling in some elections is more accurate than others? In particular, why were the polls so wrong in Alberta and BC while doing so well in Ontario during the same time period?

There are a number of explanations for why polls might fail to accurately measure public opinion. A meta-analysis of academic research investigating inaccurate polling found that explanations of polling failures centred around three explanations (Durand et al., 2010). One focuses on the

voters themselves and some change in behaviour either in terms of turnout or a change in voter preferences at the end of the campaign. Another attributes the polling failures to problems with estimation mainly inappropriate weighting or an inaccurate likely voter model. While a third deals with coverage and sampling problems including non-response bias, sample size, and coverage. A study of polling conducted during the 2004 and 2006 Canadian federal elections found that there was a systematic industry bias in 2004 and 2006 that underestimated Liberal support. Whether that bias was caused by methodological issues or a late shift in vote intentions was not clear, but the questions raised by the study are similar to those of this study (Pickup and Johnston, 2007).

In this study we adapt and test two of those explanations. We extend work on polling accuracy to provincial elections in Canada, using data from the nine provincial elections held in Canada between 2011 and 2013 to identify what might affect the accuracy of polls released in the final week of the campaign. Specifically, we attempt to empirically measure the two broad arguments above - mainly that some methodologies are better than others in measuring voter preferences and that the nature of the election itself might make accurate forecasts more difficult. We focus on the potential impact of voter turnout, differential voter turnout by age, survey methodology, the number of undecided voters and the volatility between elections on the accuracy of publicly released polls. We do not try to explain election outcomes nor do we examine the likelihood of sharp shifts in voter intention over short periods of time as may have been the case in Alberta or British Columbia.

The research question assessed in this article is: Why were the polls more accurate in some provincial elections than others?

One obvious source for error in public opinion research is the sample design and mode of data collection. Computer Assisted Telephone Interviewing (CATI) is the most traditional method for collected survey data of three used to measure voter preferences in publicly released polls. It involves a live interviewer conducting the survey over the phone, assisted by a computer program which prompts questions to the interviewer as the interview progresses (Butler, 2007: 75). The methodology for CATI surveys can vary by the type of sample it uses (random numbers vs. listed numbers), whether it includes cell phone-only households in the sample, and the call back procedures used to increase the response rate.

CATI is a better method of conducting surveys of smaller populations where the limitations of IVR (low response rate, sometimes below 10%) and internet surveys (panel size) make it difficult to conduct representative surveys of geographically targeted and small population areas. For example, it would be very difficult to conduct an IVR or internet-based survey of Prince Edward Island due to the province's small population.

Interactive-Voice Response (IVR), sometimes referred to as robo-polling, uses an automated message over the telephone to collect data. Respondents receive a call and are prompted by either a pre-recorded message or by an automated voice instructing them to use their keypad to answer questions. IVR surveys are relatively inexpensive to field and can be completed in a short period of time. The limits are that the surveys have to be much shorter than CATI or internet surveys and the respondent's identity cannot be confirmed. Moreover, most IVR polls only call a household once (industry standard is four to five times) and the field work is usually conducted in a short time frame (one or two nights). This can introduce biases into the

results because certain people may be more likely to be home and answer the phone on the evenings the poll is conducted. IVR surveys also suffer from low response rates due to the automated nature of the interview.

Unlike CATI and IVR surveys, internet surveying generally uses non-probability samples (although some panels claim to have recruited their panels using probability based sampling). In all cases, the panels consist of individuals who have agreed to participate in surveys when recruited to join the panel. Random panelists are then invited to complete the survey through email. Due to the non-probabilistic nature of survey panels and the limitations of size, representative internet surveys are more difficult to conduct in jurisdictions with small populations such as Prince Edward Island (Butler, 2007: 55). Online panels where sample for internet surveys are drawn may include anywhere between 80,000 and 500,000 Canadians. In many cases, there may not be enough sample for smaller jurisdictions to produce a representative sample large enough to make confident observations.

All three survey methodologies have advantages and disadvantages. CATI surveys are still considered by some in the industry to be the best at accurately measuring public opinion because they more closely adhere to probability theory and response rates are generally higher than IVR surveys. There is much debate over the ability of internet surveys to accurately measure public opinion. Questions frequently asked about internet surveys surround the representativeness of the panels used to conduct the research and whether the act of completing many surveys has an impact on the opinions and behaviour of panelists.

Research on comparisons between CATI and Internet surveys in the United States have found mixed results. One study found that probability sample surveys "were consistently more accurate than the non-probability sample surveys, even after post-stratification with demographics" (Yeager et al., 2011: 709) while another study refuted Yeager et al. and other studies on the basis that their findings "are based on data collected five or more years ago. The science of constructing, matching and weighting opt-in internet panels can produce data that looks remarkably like that from a landline/cell telephone survey (Ansolabehere and Scaffner, 2010: 1)."

Pollsters in Canada often debate the merits of survey methodologies. Obviously each has a self-interest in promoting the method their businesses use but the record of different methodologies has been mixed both in Canada and around the world. In the 2012 United States Presidential election, the New York Times' Nate Silver found that internet polls performed best when compared to live telephone (CATI) and IVR polls (Silver, 2012). In Canada, Eric Grenier who runs a Canadian election prediction website called ThreeHundredEight.com, has found that the record of different polling methodologies is mixed. For example, in the Ontario 2011 provincial election, he found that IVR and internet polls were the most accurate (Grenier, 2011b) while in the 2011 Canadian General Election, internet and CATI surveys produced the most accurate forecasts (Grenier, 2011a).

It is therefore unclear whether the methodology used to collect data has an impact on the accuracy of the polls. Apart from the way in which data is collected, each pollster has their own method for statistically weighting the data, dealing with undecided voters, and modeling for turnout. Nonetheless, the variation in data collection methods used in the nine provincial elections provides us with an opportunity to assess whether the methodology used impacts the accuracy of the polls. We do not expect the methodology used to collect data

to be a significant predictor of polling accuracy.

Survey methodology is not the only challenge faced by public opinion researchers. Declining voter turnout in the past three decades has meant that the population that votes is not necessarily representative as the population as a whole. If researchers aim to conduct a survey of a representative sample of the population, how do pollsters know that their sample being used to make inferences about the population is representative of the electorate?

We know, for example, from Elections Canada's voter turnout estimates that younger Canadians are far less likely to vote in federal elections. In the 2011 Canadian General Election, voters aged 18 to 29 made up approximately 20% of the eligible electorate and yet made up only 14% of the electorate that voted in the election (Elections Canada, 2011).

Changes in voter turnout and differential vote choice across demographic and regional groups make it more difficult for researchers to accurately forecast elections. If voting was compulsory as in Australia, modeling turnout would not be a problem. Declining voter turnout (see Siaroff and Wesley, and Thorlakson, this Edition) means that modeling for turnout has become a necessary but important challenge for researchers conducting polls during a provincial election campaign.

As Table 1 reports, voter turnout levels for the provincial elections held between 2011 and 2013 were mixed with a high of 76.4% in Prince Edward Island to a low of 49.2% in Ontario. Four provinces saw an increase in eligible voter turnout from the previous election (Alberta, British Columbia, Nova Scotia, and Quebec) while there was a drop in voter turnout in the other five.

Most of the time, public opinion research seeks to measure the opinions and behaviours of the entire population. Samples are drawn to be representative of the population as a whole. During an election however, and especially as election day draws near, the expectation of pollsters is that they will accurately forecast the election (in other words, measure the expected behaviour of those who will actually show up to the polls on election day).

But the tradition of trying to isolate "likely voters" is not as strong in Canada as in other countries. In the aftermath of the 2013 British Columbia elections, researchers at Ipsos-Reid and Angus Reid Strategies noted that part of the failure of polls to accurately forecast the election was due to the failure of those polls to account for different turnout rates among different demographic groups.

In fact, efforts by some pollsters to identify likely voters within their polling actually produced results that were less representative of the final election results. In British Columbia, EKOS Research (using a interactive voice response methodology) tried to identify likely voters through a battery of questions in the survey questionnaire. The result was that the vote intention of their identified "likely voters" was even less accurate than the results of its general population results (EKOS Research Associates, 2013).

We therefore expect a sharp shift in voter turnout to have a negative impact on the accuracy of polling if a pollster does not try to account for those changes in the weighting, or if the pollster does not effectively identify which respondent is actually likely to turnout on election day. If polling conducted in British Columbia is representative of that conducted in other provincial elections, then pollsters did not account effectively for turnout in their final publicly released polls.

Another challenge faced by researchers measuring voting intentions in elections is the volatility of voters (Gidengil

et al., 2012; Bischoff, 2013). Many voters resemble impulse shoppers as opposed to loyal customers to a particular brand. The weakness of party identification at the federal level has been a long established feature of the Canadian electorate going back to the 1970s (Leduc et al, 1987). But we know very little about voter loyalty at the provincial level and whether voters in Nova Scotia, for example, are more loyal to parties than in Ontario or British Columbia.

We expect that it should be more difficult to measure voter preferences in a highly volatile electorate where voters make up their minds late in an election campaign and are more likely to change preferences between elections. In other words, stable party systems are easier to measure than instable ones and the emergence of new political parties into a system further complicates attempts to measure voter preferences.

Methodology

This study analyzes the results of 36 polls released publicly in the final week of provincial election campaigns in Alberta, British Columbia, Manitoba, Newfoundland and Labrador, Nova Scotia, Ontario, Prince Edward Island, Quebec, and Saskatchewan . Data for this study was collected from news reports and the websites of public opinion research firms. The reason we use this time horizon is because not all pollsters released their final numbers one or two days before the campaign ended. We control for the number of days out from the election the poll was released within our multivariate model.

There has been little formal work in Canada on the accuracy of publicly released polls. Prior work on political polling in Canada has examined whether polls are biased against the Liberal Party in Quebec provincial elections (Durand, 2002) and how the media covers election polls in Canada (Lachepelle, 1991). The literature on polling accuracy is more developed in other countries.

In this study, polling accuracy is assessed using the absolute total error of a poll from the final outcome (popular vote) for the major political parties running in the election. We chose this measure as opposed to others due to the multi-party electoral systems in most of Canada's provinces. Other studies have used a measure developed by Martin et al (2005) (Wright et al., 2013; Panagopoulous, 2009; Callegaro and Gasperoni, 2008; and Traugott, 2005) but it only includes the accuracy of the two leading parties in an election, ignoring the accuracy of polls in measuring smaller parties. It is a measure of polling accuracy for the two leading parties in an election.

In our study, for any given poll, accuracy is defined as the total absolute error in percentage point terms. For example, if a poll forecasted that Party A would receive 50%, Party B would receive 30%, and Party C would receive 20%, but the final election results were Party A 45%, Party B 35%, and Party C 20%, the total absolute error for the poll is 10. The smaller the number, the more accurate the poll.

To explain polling accuracy, a multivariate regression model was specified using absolute error measure as the dependent variable with a number of independent variables. To test the impact of survey methodology on accuracy, we created two dummy variables for IVR and internet methodologies with CATI methodology as the comparator group. Eligible voter turnout and the absolute change in voter turnout from the previous election were used to test the impact of voter turnout while the absolute sum of the change in the proportion of vote for each of the parties from the previous election was used to measure electoral volatility[1].

The model also controls for the margin of error of the each poll and the

number of days the poll was released prior to election day.

Findings

The overall performance of public polling for provincial elections held from 2011 to 2013 was mixed. Of the nine elections, polls forecasted the final vote shares in six of the elections within the polls' margin of error while in three elections the results were much less successful. Polls conducted in the final week of the campaign in Manitoba, Nova Scotia, and Ontario were the most accurate while those in Alberta, British Columbia, and Quebec were the least accurate.

As Table 2 reports, there is a relationship between polling accuracy and change in turnout. In provinces where turnout changed substantially from the previous election, either up or down, polls were less accurate than polls conducted in provinces where voter turnout changed little from the previous election.

The relationship between electoral volatility and polling accuracy is less clear. For example, Nova Scotia, which had the second most accurate polls also had the second highest level of electoral volatility. But in Alberta, where electoral volatility was highest, the predictive accuracy of the polls in that election were lowest. Of course, the volatility in Alberta was different since it was caused by the introduction of a new party.

There was also a relationship between survey methodology and predictive accuracy of polling of the 36 final polls released during the nine provincial elections in Canada. CATI polls were generally more accurate than those conducted by IVR or internet. On average, CATI polls produced total error of 8.75 compared with 12.75 for internet polls and 13.42 for IVR polls.

When we control for survey methodology, the margin of error for each survey, the number of days between the final day of data collection and election day, absolute change in voter turnout, voter turnout, and electoral volatility, we find that only two variables had a statistically significant effect on polling accuracy.

Overall, the model explained 33.9% of the variation in polling accuracy in the nine provincial elections held in Canada from 2011 to 2013 and the largest predictor of polling accuracy of the variables included was the absolute change in voter turnout, all else being equal. A single point change in absolute voter turnout results in a 1.302 decrease in polling accuracy. The model also found that as voter turnout increased, the accuracy of polls generally improved as well (a negative relationship in the model).

While none of the other variables in the model were statistically significant, the direction and size of the b values are worth discussing. The margin of error associated with each poll released was positively related to polling error. As the margin of error of a poll increased the error of the polls increased. This finding is intuitive. Moreover, the closer a poll was released to election day, the more accurate the poll was in forecasting the election result.

The methodology used to collect survey data had little effect on the accuracy of polls released in the final week of the provincial elections, all else being equal. IVR polls were somewhat less accurate that CATI polls while there was almost no difference between internet and CATI polls in terms of effect on polling accuracy

One question the model was unable to assess was the impact of undecided voters on the final outcome of the campaign. In Alberta in particular, arguments were made that many voters made up their mind in the final weekend of the campaign, long after most pollsters were out of the field. If voters did shift their vote intentions in that final week, most pollsters were unable to capture the shift. Forum Research was the only

Table 1: Predictive Accuracy of Polling, Canadian Provincial Elections, 2011-2013

	Absolute Polling Error	Number of Polls Released in Final Week	Change in Voter Turnout	Electoral Volatility from Previous Election
PEI (2011)	7.0	1	-7.4	0.05
Manitoba (2011)	2.0	3	-1.0	0.14
Ontario (2011)	6.0	7	-3.6	0.18
Newfoundland and Labrador (2011)	12.0	2	-3.4	0.33
Saskatchewan (2011)	6.7	3	-10.0	0.27
Alberta (2012)	21.3	7	+16.4	0.54
Quebec (2012)	11.8	4	+17.2	0.29
British Columbia (2013)	17.0	6	+7.0	0.09
Nova Scotia (2013)	7.3	3	+1.1	0.39
Canada (2011)	*4.0*	*8*	*+2.3*	*0.29*

Table 2: Multivariate Model Predicting Absolute Total Polling Error

	b	Std. Error	Beta
Absolute change in voter turnout	1.302**	.377	1.036
Voter turnout	-.475*	.206	-.487
Margin of error	2.985	1.819	.323
IVR	1.667	3.624	.103
Polling dates	.376	.505	.120
Electoral volatility	-4.912	9.615	-.103
Internet	.252	3.171	.015
Constant	18.594	12.217	

r^2 = .339 ** p<0.000 *p<0.005

polling firm to be in the field over the weekend and they reported a much closer race on the eve of the election than any previous polls (Forum Research, April 22, 2012).

In order to assess the potential for late shifts in vote intention, we need data that measures the general uncertainty of the voters. In other words, we need to know if the respondents answering a survey could change their intentions before election day and to what extent. Abacus Data has such data for the last couple of provincial elections[2]. Instead of simply being asked who they would vote for, respondents ranked their likelihood to vote for each party, on a scale of 0 (completely unlikely) to 10 (would vote for sure). This question allows us to measure the level of certainty of a voter's intention and the probability that it may change before election day. While we cannot compare these likelihoods with the actual voting decision, we can compare them to the answers to the more traditional vote intention question ("which party would you vote for?"). This allows us to assess how well parties can convert likely voters into actual intended voters. For instance during the last Ontario election, 29% of respondents who rated their likelihood of voting PC as 6 on a 10-point scale also said they vote PC if the election was held at the time of the survey. In contrast, only 15% of respondents who rated their likelihood of voting NDP as 6, chose the NDP in the traditional vote intention question. This means that although the self-reported likelihood of voting for each party might be similar, the ability of a party to convert that likely voter into an actual voter may be different

depending on the party and the context of the election. Moreover, what this shows is that the electoral outcome can be quite different depending on which party is chosen by the voters who aren't absolutely certain.

Using the Ontario and Alberta data, we calculated, what we call, volatility ranges. These intervals account for the normal statistical uncertainty that arises from only having a sample of the population, but also for the fact that a lot of voters would consider voting for more than one parties. Specifically, the idea is to measure the potential impact of voters changing their minds. For instance, if there were indeed a lot of undecided in Alberta between the Wildrose Party and the PC Party, then a last minute swing of these undecided could well explain why the polls were so off.

We ran thousands of simulations where we randomized the conversion from the likelihood to vote into actual support. In simpler terms, if there are respondents torn between voting for the Liberals and the NDP, in some simulations they are allocated to the Liberals while in others, they go with the NDP. However, the simulations do not simply allocate the respondents randomly. Someone who has a likelihood to vote PC of 8 is obviously a lot more likely to end up supporting this party than someone with a likelihood of 2. The simulations account for that in a very systematic way.

In Table 3, we present these ranges for the last Ontario and Alberta elections, at 95% confidence levels.

For Ontario, it is clear the NDP had a much higher potential vote than they received in the traditional vote intention question (and also on election day). One possible explanation is that a lot of voters who were considering voting NDP and Liberal ultimately opted for the latter in order to prevent a PC victory. This type of data can help explain surprising results and tactical voting.

As for Alberta, the volatility ranges for the parties finds that either the Progressive Conservative Party or the Wildrose Party could have won the election. As opposed to the conclusions we would reach using the normal margins of error of most polls, the volatility ranges did show that both parties were virtually tied and that there could be a swing in the final days of the campaign. We also see that the ranges (or intervals) for Alberta are greater than for Ontario despite having similar sample sizes. There was indeed more undecided voters in Alberta than in Ontario. Nevertheless, the high range for the PCs is lower than the actual vote share the party received on (44%) meaning the polling error in Alberta cannot fully be explained by a last minute swing of undecided voters.

Implications

Polls measuring voter preferences in provincial elections where voter turnout changed substantially were generally less accurate than those in elections were turnout was more stable. This has important implications for polling in the future. The fact that the polls were unable to account for the sharp rise in voter turnout and what they meant for voting rates among different demographic and regional subgroups partly explains why the polls were inaccurate in the Alberta, BC, and Quebec elections.

Most pollsters already statistically weight their samples, regardless of the method the data is collected. These weights are usually there to make sure that the sample is representative of the general population, thus using census data to match the characteristics of the general population on some key variables (age, geography, gender, education and language (especially

Table 3: Volatility Ranges for Ontario 2011 and Alberta 2012

	Volatility ranges	Vote intentions (traditional question)	Actual result
Ontario 2011			
Liberal	30-40%	37%	38%
PC	29-37%	34%	36%
NDP	22-33%	24%	23%
Alberta 2012			
PC	22-40%	35%	44%
Wildrose	27-45%	39%	34%
Liberal	6-18%	10%	10%
NDP	11-20%	14%	10%

in Quebec)). Polls are less accurate when the population they are measuring does not match the sample it is using. Since many pollsters do not weight according to likely turnout, this helps explain why turnout can be such an important factor in the accuracy of pre-election polls.

According to an analysis of turnout in the 2009 provincial election by Elections BC, voters under the age of 35 only represented around 15% of the total voters in 2009. However, in 2013, pollsters were giving this group a weight closer to 30%, the actual share of people under 35 in the population. On the other hand, voters over 60 were underrepresented. Given that the BC NDP had an important lead among the under 35 but were trailing among the 60+, we can easily see how the incorrect weighting led to an overestimation of the BC NDP.

Angus Reid, chairman of Angus Reid Public Opinion concluded that an incorrect voter turnout model caused his firm's missed election call. In an op-ed in *Macleans* wrote,

> The principal flaw in our methodology was that we represented voters under 35 (where the NDP held a commanding lead) in relation to their proportionate share of the BC population (roughly 30 per cent) rather than in relation to their actual share of voters (closer to 15 per cent according to research conducted by Elections BC after the 2009 contest). Had we made this one change in our turnout projection model the final Angus Reid poll published on May 9 would have shown the NDP lead diminish to only three points (Reid, 2013:accessed online).

Similarly, a post-election analysis conducted by polling firm Ipsos Reid confirmed that modeling voter turnout was the culprit in its poor pre-election forecast in British Columbia. The firm conducted an Election Day exit survey that was weighted by age according to voter turnout levels estimated by Elections BC which accurately reflected the results of the election.

Finally, Abacus Data only conducted a poll around the mid-point of the BC election campaign (between April 23-26, so just before the debate and a little over two weeks before the election). The published poll showed the NDP ahead of the BC Liberals 43% to 33% among committed voters. However, if that data had been weighted by age according to the data from Elections BC, the NDP's lead would have been reduced to just 3-points. At that point of the campaign, it is possible that the NDP had such a lead among those likely to vote which

disappeared as election day drew near. In any case, the pollsters themselves admit that their estimates would have been more accurate had they weighted their data against likely voter turnout by age group.

We should however be cautious with the use of statistical weighting to predict the composition of the electorate. While using weights to correct our sample can help in some cases (as we just showed for BC), it does not change the fact that good research still requires good data collection and sampling strategies.

In Quebec for instance, sociologist Claire Durand (2012) found that the samples collected by polls conducted during the Quebec election had varying proportions of non-francophone. In some cases, non-francophone voters represented as much as 26%, while at other times (and most of the time), they were below 15%. The actual proportion of non-francophones in the population is approximately 20% according to the 2011 Canadian census. In itself, the under-representation of this group is not necessarily a problem. But Claire Durand showed that there was a correlation between the proportion of non-francophones and the support for the Liberals among this group. If an important group like non-francophone voters is significantly underrepresented in the original sample and therefore does not accurately reflect the groups voting intentions, then the weighting can make the forecasts even less accurate. In the Quebec case, a sample with an underrepresentation of non-francophones could have a bias against the Liberals. Simply giving more weight to these observations would clearly not eliminate this bias.

Therefore, while weighting and modeling for voter turnout can improve polling forecasts, it is not a solution for bad polling or highly unrepresentative samples. On top of that, successfully identifying likely voters is not easy. In the British Columbia election, polling firm EKOS tried to account for likely voters but its estimates among likely voters were even less accurate than results will all respondents included.

Another point to consider, beyond sampling and weighting, is the fact that measuring voting intentions is very different from measuring (say) the average age of a population. For the latter, the only uncertainty comes from the normal statistical variation (i.e: margins of error). But for the former, we also need to possibly account for fact that many voters are undecided. As we showed, this fact could have contributed to the general poor performance of the polls in Alberta. Added to the important rise in the turnout, these two factors can potentially explain the general failure of the polls for that specific election.

The expectation that polling will always accurately forecast election results is faulty and has led to the perception that pollsters should always be able to forecast the final election results. One reason for this perception has been caused by the rise of seat projection models like the one designed by one of the co-authors of this article. For seat projection models to work precisely, they require polling results to be almost perfect reflections of the final popular vote totals.

It is our hope that the news media and public recognize the role that polls can play in measuring public opinion and voter intentions while also recognizing the limitations of surveys in a volatile environment. It is incumbent upon researchers who release polls during a campaign to refine their methods and recognize the challenges that exist in trying to pinpoint the vote intention of a fickle and volatile electorate.

The results of this article also offer some lessons to those who study and conduct polling in Canadian federal elections. Polling in the past number of federal elections has been fairly accurate[3]

but there has not been a substantial shift in voter turnout since 2006 and the structure of the party system has remained relatively stable. Our research has shown that polling needs to take into account the potential for large shifts in voter participation and the impact that can have on a polls accuracy in forecasting the election results. Pollsters need to do a better job at modeling for turnout and identifying likely voters. These are lessons that only a thorough study of subnational elections, and the variation in conditions, party systems, and campaign dynamics can offer.

Finally, we cannot finish without mentioning the merits of studying provincial election polling. While academic attention to provincial polling in Canada has been scant, doing so provides many opportunities for researchers studying Canadian or comparative politics. First, there is a larger set of publicly released polls to work on than at the federal level in Canada. In the nine elections studied in this article, there were over one hundred publicly released polls. In contrast, in the 2011 federal election, 81 vote preference polls were released. Along with more data points, the polls released during subnational elections use more varied methodologies and are released from a broader set of research firms. Second, there is far more variation in party systems, voter behvaiour, and campaign dynamics across ten provincial elections than if research focuses on a single federal campaign. It may take a decade or more to observe the same range of variation in the nine provincial elections studied in this book at the federal level.

Despite these opportunities, there are some challenges when using subnational polls as data to analyze polling. In some provinces, such as Prince Edward Island, Manitoba, and Saskatchewan, there were only a few polls released during the campaign often by one or two firms. This can introduce bias as conclusions rely on only a few data points. Moreover, some firms may participate in some provincial elections but not others as was the case in the nine elections studied in this article. The inability to control for "house effect" can also impact the conclusions drawn for the analysis.

Works Cited

Andersen, Robert. 2000. Reporting Public Opinion Polls: The Media and the 1997 Canadian Election. *International Journal of Public Opinion Research* Vol. 12, 285-298.

Ansolabehere, Stephen., & Schaffner, B. F. 2011. Re-examining the validity of different survey modes for measuring public opinion in the US: Findings from a 2010 multi-mode comparison. Paper presented at the Annual Meeting of the American Association for Public Opinion Research, Phoenix.

Bischoff, Carina. 2013. "Electoral unstable by supply or demand? - an examination of the causes of electoral volatility in advanced industrial democracies." *Public Choice.* 156: 537-561.

Butler, Peter. 2007. *Polling and Public Opinion*: *A Canadian Perspective*. Toronto: University of Toronto Press.

Callegaro, Mario, & Gasperoni, G. 2008. "Accuracy of pre-election polls for the 2006 Italian parliamentary election: Too close to call." *International Journal of Public Opinion Research* .20:148-170.

Durand, Claire. 2002. "Are the polls biased against Quebec Liberals?" *Policy Options*, IRPP: March 2002.

Durand, Claire., Deslauriers, M., Goyder, J., and Foucault, M. 2010. Why do polls go wrong sometimes? Paper presented at the Annual Meeting of the American Association for Public Opinion Research, Chicago.

Durand, Claire. 2012. "Le paradoxe de l'anglo péquiste/caquiste" http://ahlessondages.blogspot.ca/2012/08/le-paradoxe-de-langlo-pequistecaquiste.html (Accessed October 27, 2013).

EKOS Research Associates. 2013. "An unapologetic analysis of the BC polling debacle" http://www.ekos.com/admin/articles/FG-2013-05-29.pdf (Accessed October 27, 2013).

Elections BC. "Voter participation by age group - May 12, 2009 Provincial General Election" http://www.elections.bc.ca/docs/stats/voter-participation-by-age-group-2009.pdf (Accessed October 27, 2013).

Elections Canada. Estimation of Voter Turnout by Age Group. http://www.elections.ca/content.aspx?section=res&dir=rec/part/estim&document=index&lang=e (Accessed October 27, 2013).

Forum Research. "Wildrose Heading for a BareMajority". 2014. http://www.forumresearch.com/forms/News%20Archives/News%20Releases/40010_Alberta_Issues_Poll_%28Forum_Research%29_%2820120422%29.pdf (Accessed August 29, 2014).

Grenier, Eric. 2011. Pollsters did better but still missed mark in Ontario". Globe and Mail Online http://www.theglobeandmail.com/news/politics/ontario-election/pollsters-did-better-but-still-missed-mark-in-ontario/article2194382/ (Accessed October 27, 2013).

Jowell, R., Hedges, B., Lynn, P., Farrant, G., & Heath, A. 1993. The 1992 British Election: The Failure of the Polls. In *Public Opinion Quarterly*. 57: 238-263.

Lachapelle, G. 1991. *Polls and the Media in Canadian Elections: Taking the Pulse* Study of the Royal Commission on Electoral Reform and Party Financing. Toronto: Dundurn Press.

LeDuc, L. 1987. "Partisan instability in Canada: Evidence from a new panel study" *American Political Science Review* 78:470-484.

Martin, Elizabeth, Traugott, M.W., & Kennedy, C. 2005. "A review and proposal for a new measure of poll accuracy." *Public Opinion Quarterly* 69: 342-369.

Mitofsky, Warren. 1998. "Was 1996 a worse year for polls than 1948?" *Public Opinion Quarterly*. 62: 230-249.

Gidengil, Elisabeth., Nevitte, N., Blais, A., Everitt, J., and Fournier, P. 2012. Dominance and Decline: *Making Sense of Recent Canadian Elections*. University of Toronto Press.

Panagopoulos, Costas. 2009. "Preelection poll accuracy in the 2008 general elections." *Presidential Studies Quarterly*. 39: 896-907.

Pickup, Mark. and Johnston, R. 2007. Campaign trial heats as electoral information: Evidence from the 2004 and 2006 Canadian federal elections., *Electoral Studies*, 26:460-476.

Reid, Angus. 2013. "What went wrong with the polls in British Columbia?" *Macleans* http://www.macleans.ca/news/canada/angus-reid-what-went-wrong-with-the-polls-in-british-columbia/ (Accessed October 27, 2013)

Silver, Nate. 2012. "Which Polls Fared Best (and Worst) in the 2012 Presidential Race" *New York Times Online* http://fivethirtyeight.blogs.nytimes.com/2012/11/10/which-polls-fared-best-and-worst-in-the-2012-presidential-race/ (Accessed October 27, 2013).

Traugott, Michael. 2005. "The accuracy of the national preelection polls in the 2004 presidential election campaign." *Public Opinion Quarterly* 69: 642-654.

Wright, Malcolm., Farrar D., and Russell, D. 2013. "Polling Accuracy in a Multiparty Election." *International Journal of Public Opinion Research* 25:.

Yeager, David, Krosnick, J., Chang, L., Javitz, H., Levendusky, M., Simpser, A., & Wang, R. "Comparing the Accuracy of RDD Telephone Surveys and Internet Surveys Conducted with Probability and Non-Probability Samples." *Public Opinion Quarterly* 74:709-747.

Endnotes

[1] For example, in Alberta, the electoral volatility score is .54 calculated by adding the change in each party's vote share from the 2009 election = PC abs(53-44) + WR abs(7-34) + Liberal abs(26-10) + NDP abs(9-10) + Other abs(0-1) = 54

[2] But not for all, thus not permitting us to add a variable measuring this general uncertainty in our previous regression

[3] Using the same measure of predictive accuracy, the accuracy scores for the last three federal elections were: 0.140 (2011), 0.121 (2008), 0.108 (2006). Polling in the final week of the campaign was more accurate, on average, than six of the nine provincial elections studied in this article.